Section 2.

DIGITAL-SPECIFIC TEST RESULTS

Definitions:

TOV = Threshold of Visibility
POU = Point of Unusability
POA = Point of Acquisition

2.1. Free Form Viewing

a. Conditions:

Images Viewed:

Stills:	*Table & Chairs, Vines, Wavy Wall, Tulips, Sculptures, Fruits & Vegetables, Toys, Girls with Toys, Memorial Arch, Woman with Roses, Lorain Harbor, Flower on Plate*
60 f/sec film:	*Amusement Park, Turbo, Skiers, Schoolyard, Helicopter*
24 f/sec film:	*Stairway, Mirror, Christa*
30 f/sec film:	*Bridge, Fountain*
Motion Video:	*Window, FAX Machine, Mannequins, Living Room, Den, Park Ride, Audience, Woman & Room, Lamp, Texas Dude, Crosswalk, Ax Murderer, Buckingham Palace, Snow Trees, End Zone, Dream Team, Golf, Roller Coaster #1, Advisory Committee (Ducks), Picnic with Ants, Reflections, Skull*
Computer-generated:	*Tube, Slinky, Rotating Pyramids, Cheshire Cat, Clock #1, Connections*

Explanation of Test:

We viewed all the above images in both 1080-I and 720-P formats. For each image, the five observers compared (in this order) the image at the input to the GA hardware, the image at the output of the GA hardware as rendered in the same format as the input, and the image at the output of the hardware as rendered after conversion by the GA format converter to the other format.

b. Commentary on 1080-I Format:

Observations:

The following Table 2-1 shows the average scores assigned by the observers to each image. The scoring is a 10-point scale, where 10 is best and 1 is worst. The scale is deliberately different from a CCIR scale, since CCIR voting procedures were not employed (*e.g.*, the observers knew whether they were seeing input, output, or format-converted images).

Table 2-1

	1080 - I		
Test Image	**Reference Input 1035I**	**Output 1080I**	**Output Scan Converted 720P**
Vines	9.2	9.0	9.0
Wavy Wall	9.5	9.5	9.5
Woman with Roses	9.5	9.4	9.3
Lorain Harbor	9.5	9.4	9.4
Flower on Plate	9.7	9.7	9.6
Park Ride	8.7	7.6	7.8
Audience	8.9	8.6	8.3
Lamp	8.5	7.8	7.9
Texas Dude	9.0	8.6	8.5
Bridge	7.8	7.5	7.5
Helicopter	8.8	8.6	8.3
Skiers	9.0	8.5	8.5
School Yard	8.6	7.9	8.1
Amusement Park	8.8	8.2	8.2
Turbo	8.4	8.1	8.1
Slinky	9.4	9.3	9.3
Tube	9.5	9.2	9.1
Stairway	7.5	7.4	7.4
Reflections	9.3	9.3	9.3
Skull	9.2	7.3	6.4
Metal Table & Chairs	9.3	9.0	8.8
Tulips	9.7	9.3	9.3
Sculptures	9.6	9.3	9.3
Fruit & Vegetables	9.4	9.3	9.1
Toys	9.0	8.9	8.3
Girl with Toys	9.5	9.2	9.1
Memorial Arch	9.4	9.3	9.0
Cheshire Cat	9.4	9.4	9.3
Rotating Pyramids	9.2	8.1	8.1
Crosswalk	9.1	8.3	8.0

Table 2-1 (continued)

	1080 - I		
Test Image	**Reference Input 1035I**	**Output 1080I**	**Output Scan Converted 720P**
Ax Murderer	9.1	9.0	9.0
Buckingham Palace	9.4	8.9	8.9
Snow Trees	9.4	8.5	8.5
End Zone	8.9	8.5	8.6
Dream Team	9.0	7.5	7.3
Golf	9.2	9.0	9.0
Roller Coaster #1	9.4	8.6	8.5
Advisory Committee (Ducks)	9.2	8.5	8.4
Mirror	9.1	8.9	8.9
Christa	7.8	7.6	7.6
Fountain	8.9	8.6	8.6
Clock #1	9.0	8.3	8.3
Connections	9.5	9.4	9.2
Picnic with Ants	9.5	6.8	6.7
Window	9.2	8.6	8.7
FAX Machine	9.2	8.5	8.3
Mannequins	8.9	7.9	8.0
Living Room	8.9	7.8	7.7
Den	9.1	8.6	8.2
Woman & Room	9.4	8.4	8.2

Summary Observations and Judgments:

In general, the images were rendered quite well. It is the judgment of these observers that image quality was clearly better than that of any of the previous systems. The level of the usual compression artifacts (quantization noise and blockiness) was lower than we had observed on previous systems. We also did not observe even the low levels of regular "pulsing" of the quantization noise that we had seen previously.

Despite the strong endorsement of the GA system's image quality in the previous paragraph, we were able to find compression artifacts on some of the highly stressful images. The most stressful of these images were new motion sequences used for the first time in this round of tests. Examples:

Dream Team: The combination of extremely rapid motion and flash bulbs (which caused big changes in image brightness in large areas of the picture and

in so doing rendered predictive coding extremely difficult) caused noticeable blockiness to appear in the image.

Skull: The strobe light effect caused blockiness to become visible for the same reason as the flash bulbs in *Dream Team.*

Crosswalk, Advisory Committee, Buckingham Palace, Roller Coaster: Observers noticed a slight loss of detail or resolution in the output image. *Crosswalk* had slight quantization noise in some highly detailed still areas of the picture (*e.g.*, manhole cover).

Saturated reds had a tendency to show compression noise more readily than other portions of images (example: *Rotating Pyramids).*

In general, the quality of the film source material, especially at 24 frame/sec, was a more serious limitation of the image quality than any of the compression artifacts. The primary limitation was visible film grain and noise. The input images from film were poorer quality than the output images from most other sources. An arguable exception to this statement was the most stressful portion of *Dream Team.*

Motion judder and jagginess on the edges of the computer-graphic input images were also artifacts judged comparable in importance to the level of compression artifacts induced by the GA system.

Interlace artifacts on video-sourced images seemed less important limiters of image quality than the film quality limitations. Comparison of the degree of annoyance of interlace artifacts with the annoyance of computer-graphic judder is a matter of taste, and we offer no comment.

In general, the image quality after conversion to the other format (in this case, to 720-P) was judged slightly poorer than the image rendition in the "intended" format. The quality loss was generally manifested as a slight loss of resolution and a slight increase in noise.

c. Commentary on 720-P Format:

Observations:

Table 2-2 shows the average scores assigned by the observers to each image. The scoring is a 10-point scale, where 10 is best and 1 is worst. The scale is deliberately different from a CCIR scale, since CCIR voting procedures were not employed (*e.g.*, the observers knew whether they were seeing input, output, or format-converted images).

Table 2-2

	720 - P		
Test Image	**Reference Input 720P**	**Output 720P**	**Output Scan Converted 1080I**
Vines	9.2	9.1	8.6
Wavy Wall	9.2	9.1	8.9
Woman with Roses	9.2	9.0	8.7
Lorain Harbor	9.3	9.1	8.9

Table 2-2 (continued)

	720 - P		
Test Image	**Reference Input 720P**	**Output 720P**	**Output Scan Converted 1080I**
Flower on Plate	9.4	9.3	9.2
Park Ride	8.8	8.3	7.8
Audience	8.8	7.9	7.3
Lamp	8.8	8.4	8.1
Texas Dude	8.6	8.3	7.9
Bridge	8.0	7.8	7.3
Helicopter	8.5	7.6	7.1
Skiers	8.9	8.8	8.5
School Yard	8.7	7.7	7.6
Amusement Park	8.3	8.0	7.6
Turbo	8.5	8.0	8.0
Slinky	9.6	9.4	8.7
Tube	9.4	9.4	8.8
Stairway	8.0	7.8	7.6
Reflections	9.3	9.2	9.2
Skull	8.7	7.0	6.4
Metal Table & Chairs	9.2	8.7	8.4
Tulips	9.4	8.5	8.4
Sculptures	9.3	8.5	8.4
Fruit & Vegetables	9.5	9.1	8.8
Toys	9.1	8.8	8.5
Girl with Toys	9.0	9.0	9,0
Memorial Arch	9.5	9.4	9.1
Cheshire Cat	9.1	8.9	8.6
Rotating Pyramids	8.7	8.2	7.7
Crosswalk	9.1	8.5	8.3
Ax Murderer	9.1	9.0	8.6
Buckingham Palace	9.2	9.1	9.1
Snow Trees	9.3	9.0	9.0
End Zone	9.1	8.6	8.3
Dream Team	9.2	8.6	8.6

Table 2-2 (continued)

Test Image	Reference Input 720P	Output 720P	Output Scan Converted 1080I
Golf	9.2	9.1	9.1
Roller Coaster #1	9.4	9.0	8.8
Advisory Committee (Ducks)	9.2	8.7	8.6
Mirror	8.4	8.4	8.2
Christa	8.2	8.0	8.0
Fountain	8.6	8.6	8.6
Clock #1	8.8	8.6	8.5
Connections	9.3	9.0	9.0
Picnic with Ants	9.5	6.7	6.2
Window *Camera Generated*	8.4	7.4	6.8
Window *Transconverted*	9.0	8.7	8.6
FAX Machine *Camera Generated*	8.9	7.8	7.5
FAX Machine *Transconverted*	8.8	8.6	8.2
Mannequins *Camera Generated*	8.4	7.3	6.9
Mannequins *Transconverted*	8.5	8.0	8.0
Living Room *Camera Generated*	8.4	7.3	6.7
Living Room *Transconverted*	8.8	7.9	7.7
Den *Camera Generated*	8.2	7.4	7.1
Den *Transconverted*	8.7	8.4	8.2
Woman & Room *Camera Generated*	8.8	8.3	7.8
Woman & Room *Transconverted*	9.1	8.9	8.3

Summary Observations and Judgments:

In general, the images were rendered quite well. Image quality was comparable to that of the 1080-I system, and most of the observations and judgments about 1080-I apply to 720-P as well. It is the judgment of these observers that image quality of the 720-P system was, like that of the 1080-I system, clearly better than that of any of the previous systems. The level of the usual compression artifacts (quantization noise and blockiness) was lower than we had observed on previous systems. We also did not observe even the low levels of regular "pulsing" of the quantization noise that we had seen previously.

Despite the strong endorsement of the GA system's image quality in the previous paragraph, we were able to find compression artifacts on some of the highly stressful images. Examples:

Skull: The strobe light effect caused blockiness to become visible.

Rotating Pyramids: Quantization noise was visible. The observers also felt that this image and some others as well *(Dream Team)* were softer (lower resolution) in their 720-P source form than in the 1080-I source form.

Audience, Helicopter: Some quantization noise was visible.

We noted that *Dream Team* (probably the most difficult sequence in 1080-I) was rendered better in 720-P. There was less blockiness in the rapid motion.

Saturated reds had a tendency to show compression noise more readily than other portions of images.

In general, the quality of the film source material, especially at 24 frame/sec, was a more serious limitation of the image quality than any of the compression artifacts. The primary limitation was visible film grain and noise. The input images from film were poorer quality than the output images from most other sources.

Motion judder (e.g., *Clock, Connections*) and jagginess on the edges (e.g., *Clock*) of the computer-graphic input images were also artifacts judged comparable in importance to the level of compression artifacts induced by the GA system.

In general, the image quality after conversion to the other format (in this case, to 1080-I) was judged slightly poorer than the image rendition in the "intended" format. The quality loss was generally manifested as a slight loss of resolution and a slight increase in noise.

2.2. & 2.8. Scene Cuts and Video Coder Overload

a. Isolated Scene Cuts

Conditions:

The sequences observed were:

Cuts among all combinations of *Turbo, Park Ride, Lorain Harbor.*
Cuts among all combinations of *Rotating Pyramids, FAX Machine, Sculptures.*
Cut to and from *Skiers* and *Zone Plate.*

Cut to and from *Schoolyard* and *Table & Chairs*
Cut to and from *Dream Team* and *Table & Chairs*

Observations for 1080-I:

In general, scene cuts were rendered well; it is the judgment of these observers that scene cuts were rendered better in this system than in any of the original systems. Transient effects were noticeable in real time only for the cut from *Dream Team* into the complex still *Table & Chairs*. The effect was gone very quickly, so quickly that it defied accurate characterization; it seemed to be some combination of quantization and blockiness.

We also observed all the cuts in slow motion. Only *Schoolyard* into *Table & Chairs* and *Dream Team* into *Table & Chairs* showed any effects visible in slow motion, and these effects were a slight blockiness.

Observations for 720-P:

In general, scene cuts were rendered well; it is the judgment of these observers that scene cuts were rendered better in this system than in any of the original systems. Noticeable transient effects were visible in real time only for the cut from *Dream Team* into the complex still *Table & Chairs;* very slight effects were noticeable in real time for the cut from *Schoolyard* into *Table & Chairs*. The effects were gone very quickly, so quickly that they defied accurate characterization; they seemed to be some combination of quantization and blockiness.

Comments on the GA system in both 1080-I and 720-P modes:

Scene cut performance was so remarkably better than any of the previous systems that we questioned the GA about the reasons. It was stated that the coding algorithm now specifically recognized scene cuts and scheduled an I-frame at the first frame of the cut.

b. Non-Isolated Scene Cuts

Conditions:

The observed sequences were "Interferer," "Mixed Adjacencies," and "Flash Frames."

Observations for 1080-I:

The "Interferer" and "Mixed Adjacencies" sequences showed no transient effects visible in real time.

A slight blockiness was visible in the *Eye Chart* portion of "Flash Frames" in real time.

We judge performance of this GA system in this test better than any of the previous systems.

Observations for 720-P:

"Interferer" showed slight blockiness. The "Mixed Adjacencies" and "Flash Frames" sequences showed no transient effects visible in real time.

We judge performance of this GA system in this test better than any of the previous systems.

2.3. Threshold Characteristics

The observers used four images that spanned a wide range of "difficulty" in order to determine if the threshold performance varied with the complexity of the image. The images were: *Lamp* (easy motion), *Texas Dude* (moderate motion), *Rotating Pyramids* (complex motion), and *Woman with Roses* (still).

Test Results:

a. Random Noise - Effects on Video:

Conditions:

Images = *Lamp, Texas Dude, Rotating Pyramids, Woman with Roses*
Desired level = strong (-38 dBm)

The observers examined the images over a range of C/N values that covered the entire threshold. Comments are recorded. We also verified the BER measurement technique; the D/U corresponding to the TOV as measured by viewing the image compared sufficiently closely with the D/U corresponding to a BER of 3×10^{-6}.

C/N (dB)	Comments
Lamp:	
14.91	TOV.
14.41	POU.
Texas Dude:	
15.16	TOV.
14.41	POU.
Rotating Pyramids:	
15.16	TOV.
14.41	POU. Observers noted block errors, displaced slices, slice errors, freezing, "chicken wire" arrays of block outlines, and panel boundaries.

Woman with Roses:

14.66	TOV.
13.91	POU. (POU definition is arbitrary on a still.)

b. Random Noise - Effects on Audio:

We tested 4 levels of random noise impairment, ranging from the TOV to the POU of *Texas Dude*. The 5.1-channel audio was short snippets as used for other audio tests. The duration of the audio snippets was not as long as would have been desired for this test; comments are recorded below, subject to this caveat. Audio did not fail before video.

D/U (dB)	Comments
15.29 (TOV)	No audio impairments. No noticeable video impairments.
15.04	No audio impairments. Visible video impairments.
14.79	Audio drop-outs judged "quite annoying." Video contained severe blockiness and freezing.
14.54 (POU & POF)	Audio completely muted. Video not usable.

c. Impulse Noise:

The test checked the various performance thresholds for various images subjected to impulse noise impairment.

For the *Lamp* image, the tests were performed on two different days, with somewhat different results. The first day's test was intended to verify the BER measurement technique; it was agreed that the D/U corresponding to the TOV as measured by viewing the image compared sufficiently closely with the D/U corresponding to a BER of 3×10^{-6} that the BER measurement technique was validated for these purposes.

Lamp:

TOV: D/U = +0.38 dB on day of BER validation
= -0.48 dB on day of image comparisons

Threshold	D/U (dB)	Comments
TOV	-0.48	Small areas of noise bursts, some freezing
POU	-2.48	
POF	-2.98	Noise bursts; visible blocks, macroblocks, & slices; image freezes for long time

Texas Dude:

Threshold	D/U (dB)	Comments
TOV	+0.02	Small areas of noise bursts, some freezing
POU	-2.02	Blocks, slices, "glass tiles"
POF	-2.52	Long freezes

Rotating Pyramids:

Threshold	D/U (dB)	Comments
TOV	+0.23	Same comments as above
POU	-1.77	
POF	-2.52	

Woman with Roses:

Threshold	D/U (dB)	Comments
TOV	-0.77	
POU	-3.77	
POF	-4.77	POU & POF were very much a judgment call for this still image. Concealment served to hide the impairments. It was clear that the image updated only very slowly, but even slow updates are effective on a still image.

2.4. Susceptibility to Random Noise in Video Source

Test Conditions

Images: *Lorain Harbor, Rotating Pyramids, Zone Plate*

Maximum available source noise level = 140 mVrms into each of RGB

For impaired channel, desired signal level = strong (-28 dBm)

Test Results: 1080-I Format

a. Unimpaired Channel:

Comments:

The purpose of this test was to determine whether the coding algorithms exacerbated noise in the video source. The system performed well in this test. Coding artifacts were visible only at or near the maximum levels of source noise available from the test bed.

Specific Observations:

Lorain Harbor:

The "just perceptible" level of source noise, judged on the video image output from the GA hardware, was with a source noise 18 dB below the 140 mVrms maximum available. At this "just perceptible" level, the codec did not enhance the noise.

At the maximum available noise (140 mVrms), there was an increase in the blockiness of the image. (We were instructed to comment specifically on a comparison with the POU; the image, even with maximum available noise was much better quality than POU.)

Rotating Pyramids:

The "just perceptible" level of source noise, judged on the video image output from the GA hardware, was with a source noise 18 dB below the 140 mVrms maximum available. At this "just perceptible" level, the codec did not enhance the noise.

At the maximum available noise (140 mVrms), there was an increase in the blockiness of the image. (We were instructed to comment specifically on a comparison with the POU; the image, even with maximum available noise was much better quality than POU.)

Zone Plate:

The zone plate was set to include high frequencies in the sweep. At these settings, coding artifacts were visible even without added source noise. Artifacts included quantization noise and blockiness, and also a sort of "boiling" of the image in high frequency regions. These artifacts made it difficult to perform this experiment. By switching the peaking filter on and off, we noted that it caused additional visible artifacts in the form of alias spectra in this image. The various coding artifacts masked the noise and vice versa. We judged the "just perceptible" noise level to be 15 dB below the 140 mVrms maximum available.

b. Impaired Channel:

Comments:

The summary of the test results is that source and channel noise are independent.

Test Procedure:

C/N values were set as listed below. The level of source noise was adjusted to the 140 mV maximum and image behavior noted.

Specific Observations

Rotating Pyramids:

C/N (dB)	Observations
20	Same performance as unimpaired channel
18	Same performance as unimpaired channel
16	Same performance as unimpaired channel
15	Approximate TOV.

Test Results: 720-P Format

a. Unimpaired Channel:

Comments:

The purpose of this test was to determine whether the coding algorithms exacerbated noise in the video source. The system performed well in this test. The addition of source noise caused only very slight increase in image artifacts.

Specific Observations:

Lorain Harbor:

The "just perceptible" level of source noise, judged on the video image output from the GA hardware, was with a source noise 21 dB below the 140 mVrms maximum available. At this "just perceptible" level, there was a slight pulsing character to the noise, which was judged to be an artifact of the coding.

At the maximum available noise (140 mVrms), there was an increase in the blockiness of the image. (We were instructed to comment specifically on a comparison with the POU; the image, even with maximum available noise was much better quality than POU.)

Rotating Pyramids:

The "just perceptible" level of source noise, judged on the video image output from the GA hardware, was with a source noise 21 dB below the 140 mVrms maximum available. At this "just perceptible" level, slight image blockiness was visible.

At the maximum available noise (140 mVrms), there was an increase in the blockiness of the image. (We were instructed to comment specifically on a comparison with the POU; the image, even with maximum available noise was much better quality than POU.)

Zone Plate:

The zone plate was set to include high frequencies in the sweep. At these settings, coding artifacts were visible even without added source noise. Artifacts included quantization noise and blockiness, and also a sort of "boiling" of the image in high

frequency regions. These artifacts made it difficult to perform this experiment. We noted that the peaking filter caused additional visible artifacts in the form of alias spectra in this image. The various coding artifacts masked the noise and vice versa. We judged the "just perceptible" noise level to be 15 dB below the 140 mVrms maximum available.

b. Impaired Channel:

Comments:

The summary of the test results is that source and channel noise are independent.

Test Procedure:

C/N values were set as listed below. The level of source noise was adjusted to the 140 mV maximum and image behavior noted.

Specific Observations:

Rotating Pyramids:

C/N (dB)	Observations
20	Same performance as unimpaired channel
18	Same performance as unimpaired channel
16	Same performance as unimpaired channel
15.2	Approximate TOV
15	Worse than TOV

2.5. Motion-Compensation Overload

Conditions:

Image = *Girls with Toys*
Pixar-generated motion, measured in picture heights per second (ph/sec.). This method of motion generation has no temporal filtering.

Test Results: 1080-I Format

Horizontal Motion:

Speed	Comments
0.2 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.
0.4 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.
0.6 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.

0.8 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.

Vertical Motion:

Speed	Comments
0.2 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.
0.4 ph/sec.	More visible quantization noise.
0.6 ph/sec.	Quantization noise even more visible. Some blockiness.
0.8 ph/sec.	More coarseness in quantization and somewhat increased blockiness.

Diagonal Motion:
(This was specified as 0.n ph/sec. in both X and Y directions.)

Speed	Comments
0.2 ph/sec.	Visible quantization noise, especially in red areas.
0.4 ph/sec.	Noise in red areas more visible.
0.6 ph/sec.	Visible quantization noise and blockiness, especially in red areas.
0.8 ph/sec.	More serious noise and blockiness.

Test Results: 720-P Format

Horizontal Motion:

Speed	Comments
0.2 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.
0.4 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.
0.6 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.
0.8 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.

Vertical Motion:

Speed	Comments
0.2 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.
0.4 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.
0.6 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.
0.8 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.

Diagonal Motion:
(This was specified as 0.n ph/sec. in both X and Y directions.)

Speed	Comments
0.2 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.
0.4 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.
0.6 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.
0.8 ph/sec.	No noticeable artifacts beyond the usual slight quantization noise. No motion compensation overload effects.

Observer Comments:

These results show clearly that the system performed better in 720-P mode than in 1080-I mode. The observers do not understand this result, since this test should be a simple examination of motion vector range. The observers asked the following questions of the Grand Alliance: 1) Is the search range in the vertical direction greater in the 720-p mode than in 1080-I mode? 2) Is there some problem in the GA hardware when it operates in 1080-I mode? 3) Is there some kind of aliasing from lack of vertical-temporal filtering in the PIXAR source images?

Grand Alliance Comments:

In the prototype encoder, the horizontal (H) search range is ±127.5 pels for P-frames and ±63.5 pels for B-frames and the vertical (V) search range is ±31.5 pels for P-frames and ±31.5 pels for B-frames. These ranges are used in both 1080-I and 720-P modes, but the 1080-I mode uses one B-frame for each I- or P-frame while the 720-P mode uses two B-frames for each I- or P-frame. Taking the P-frame case as the limiting one, the ranges for 1080-I are 1 pw/sec H and 0.4375 ph/sec V, while those for 720-P are

2 pw/sec H and 0.875 ph/sec V. These, of course, are implementation issues and not necessarily indicative of what will happen in the future.

2.6. Multiple Impairments: Noise and Co-Channel NTSC into ATV

This test was perform using the BER technique and is reported under the ATTC test results.

2.7 Time Varying Channel Impairments

Conditions:

Image = *Crosswalk*

Test Results and Commentary:

For this test, a time-varying signal is generated by summing the desired signal with a 0.1-Hz-offset version of itself. The offset signal is added to the main signal at increasing levels. Observers comment on the observed impairments and on recovery of the system during portions of the 10-second period when the signal is strong. This shows how the system "coasts" through and recovers from signal amplitude reductions of various sizes.

The starting point was a fairly small added-signal impairment (~ 20 dB below the desired signal) in addition to the added noise. The desired signal was set at the strong level (-38 dBm); the C/N was set to ~ 16 dB, which is about 1 dB above the TOV. Thus, the added offset signal caused the amplitude to fluctuate around the threshold; as the impairment level was increased, observers recorded the results below.

a. Results for 1080-I System:

Added Signal Level (dB) below desired	Comments
-19.1	No effects.
-17.1	No effects.
-15.1	Short bursts of blockiness during "nulls." Rapid recovery.
-14.1	Long freezes and highly visible blocks and slices. Rapid recovery.
-13.1	Long freezes and highly visible blocks and slices. Rapid recovery.

b. Results for 720-P System:

Added Signal Level (dB) below desired	Comments
-19.1	No effects.
-17.1	No effects.
-15.1	No effects.
-14.1	Long freezes and highly visible blocks and slices. Rapid recovery.
-13.1	Long freezes and highly visible blocks and slices. Rapid recovery.

2.9. Film Mode

Test Conditions:

Desired level = "strong"
Switch input back and forth between 24 frame/sec film (*Christa*) and 60 field/sec video (*Rotating Pyramids*); GA system hardware switched modes automatically.
Test performed in 1080-I mode only because automatic switch is not implemented in 720-P hardware

Test Results:

Some blockiness was visible on *Rotating Pyramids* after the mode switch. The blockiness was more apparent when scrolling text was present. The blocks were not apparent immediately; they became apparent after < 1 second disappeared in an additional ~ 1 sec. The reason for the delay is not clear. The observers understood that the GA system, as a precaution against premature decisions, delays transitions into the film mode, but we did not believe such delay is employed for transitions into video.

Grand Alliance Comments:

The prototype encoder can invoke Film Mode only for groups of 5 frames. Accordingly, the transition from Film Mode to Video Mode may not coincide with the cut. Further, due to an implementation constraint in the decoder, the encoder has to insert an I-frame at the transition (we have notified the Advisory Committee of this). We suspect that what you [observers] are seeing is a transient effect in rate control due to the close proximity of 2 I-frames, one consequential to the cut and one flagging the transition to Video Mode (for benefit of the prototype decoder). If the latter were not required, the effect might have been less pronounced.

No artifacts were visible for transitions into the film mode.

2.10. Video Quality/Auxiliary Data Tradeoff

Various images were examined with the auxiliary data rate set at 256 Kb/sec, 1 Mb/sec, 2 Mb/sec, 3 Mb/sec, and 4 Mb/sec. The image data rate was reduced correspondingly.

Results:

We found the performance of the system to depend on scene content. Most scenes showed little or no artifacts at the full video rate and little or no increase in artifacts as the auxiliary data rate was increased to 3 Mbit/sec. At the 4 Mbit/sec data rate, however, the more challenging of these scenes showed a clear increase in the visibility of artifacts. One very challenging scene exhibited slight artifacts at the full video rate. For this scene, performance visibly deteriorated as the auxiliary data rate was increased.

Examples of the effects noticed at 4 Mbit/sec.:

In the *Toy* still, a just noticeable increase in noise in highly saturated chroma was noticed, along with slight blurring of details (e.g., green Lego bricks).

In the *Rotating Pyramids* and *Den* sequences, a just noticeable increase in noise on moving objects and in blocking was noticed (Noise increased in the man's sport coat in *Den* and in highly saturated chroma in *Pyramids*. Marble-patterned base in *Pyramids* exhibited blocking. *Pyramid* degradation was noticed with scrolling text).

Because the nature of the degradation seemed to be more visible during complex motion, the Expert Group examined *Dream Team* (considered to be more challenging but still realistic), with the following observations:

At zero and 256 Kb/sec auxiliary data rate, just noticeable blocking was seen (on "lay-up" portion of sequence).

From 1 Mb/sec to 4 Mb/sec the amount of blocking increased from noticeable to annoying. The blocking also was noticeable on other (less stressful) portions of the sequence (such as running down court) as the auxiliary data rate was increased.

Comment: Care must be exercised in combining an auxiliary channel with a high data rate together with video scenes with high peak complex motion; subjective degradation of the video may increase rapidly as channel capacity is diverted from video to auxiliary data.

2.11. Effects of Concatenation

a. Effects on Video:

Various video source material was passed through the GA system twice. (The output from the first pass was recorded, and then this recorded signal was passed through the GA system again.)

Results for 1080-I:

Somewhat more noise was visible after concatenation (e.g., *Skiers*). Sometimes the noise pulsed (e.g., *Table & Chairs*). Significantly more blockiness was visible on *Dream Team*. Observers judged image quality to be acceptable (marginally) as High Definition after concatenation, with the arguable exception of *Dream Team*.

Results for 720-P:

More blockiness and noise were visible after concatenation. These effects were worse with 720-P than 1080-I. There were distinctly visible degradations after concatenation on essentially all images. Noise source material (e.g., *Helicopter, Woman in Room*) showed bigger degradations after concatenation.

b. Effects on Video with Horizontally Scrolling Characters:

This test was applied to 1080-I only.

Scrolling characters were overlaid onto the image after its first pass through the GA system. The image with the overlaid characters was then passed through the GA system again. There were no special deleterious effects caused by the scrolling text (a possible exception is a slight worsening of *Dream Team,* as reported by one of the observers).

In a separate test, a source image with scrolling text was passed through the GA system twice (the text itself was concatenated). The text itself was poorer (somewhat irregular edges) after concatenation, but no other effects were observed.

c. Grand Alliance Comments:

We believe that, given analog timing variability, it is unlikely that the block boundaries were aligned from pass to pass. Accordingly, performance likely was worse than would be encountered in a studio. It is also possible that peaking, which was higher in 720-P than in 1080-I, contributed to the 2-pass artifacts.

2.12. 1035-to-1080 Transconverter Tests

The purpose of the test was to examine the transconverter for suitability as a source of 1080-line video that could be derived from 1035-line cameras and tape recorders. Observers compared input images, images at the output of the NHK transconverter, and images at the output of the GA system hardware.

Images were: *Table & Chairs, Ducks, Crosswalk, Rotating Pyramids, Cheshire Cat.*

Test Results:

Image at output of transconverter:

Most images were judged slightly softer than the original source, especially *Ducks* and *Rotating Pyramids*. *Rotating Pyramids* "shimmered" slightly in red areas.

Image at output of GA hardware:

The images were almost indistinguishable from those at the output of the transconverter. Only the usual very slight compression artifacts were visible (example: slight quantization noise visible in *Crosswalk.*) The transconverter did not impact the performance of the video coder.

2.13. Long-Form Viewing

Test Conditions:

The test was performed at 1080-I only, due to lack of source material at 720-P.

On all tests, random RF noise impairment was added at a level 1 dB below TOV.

Results:

Video Source Material with 2-channel Audio:

There were no artifacts attributable to the random RF noise. Image quality was quite good. There were no unusual or unexpected artifacts; nothing new was revealed by the "long" viewing experience.

Observers noted low-level quantization noise in some complex scenes (those including the bricks and the cobblestones). Observers also noted image blockiness in the basketball sequence from which the *Dream Team* test image was extracted (this was also noted in the Free-Form Viewing test).

Observers noted two audio drop-outs, but they were determined to be present in the source, so they were not a GA-induced artifact.

Film Source Material *(Hunt for Red October)* without Audio:

There were no visible artifacts caused by the GA system. Observers noticed some "dirty window" effects, but they were present in the source material and were not worsened by the coder.

2.14. Long-Form Film with Audio

Test Conditions:

Test conducted using 1080-I system in film mode. Material was put through hardware and viewed in real time, not on tape.

Test material was excerpts from *Hunt for Red October*.

Test conducted with random noise impairment constantly present, at a level of D/U = 16.17 dB, which was determined to be about 1 dB better signal conditions than TOV.

Results:

There were no artifacts attributable to the random RF noise. Image quality was quite good. There were no unusual or unexpected artifacts; nothing new was revealed by the "long" viewing experience. Observers noticed some "dirty window" effects, but they were present in the source material.

There were no audio artifacts attributable to the random RF noise.

2.15. Live Camera Scenes

The purpose of the test was to use an actual 1080-line camera as the video source. (For other tests, the images were derived from 1035-line cameras.)

Although several potential problems were noted in the output image, examination of source exhibited the same problem.

No image degradation was observed when the GA system processed these live camera images.

2.16. Comparative Assessments of Unimpaired Image Quality

Test Procedure:

The purpose of the test was to compare tapes of each of the previous digital systems with the present GA system in both 1080-I and 720-P modes. The same images were used for this comparison; because some of the images listed in the Free-Form Viewing Test of the GA system were new for this round of testing, they are not included in the comparisons below.

The observers viewed tapes prepared in advance by the ATTC. Each image or sequence was viewed in rapid succession on all systems before moving on to the next image or sequence. The "reference" was the 1035-I source.

Note that, for this test, the 1080-I and 720-P GA formats had different input sources for the motion video material. The 720-P camera was used as source for these sequences so that the results could be compared with the earlier progressive systems (DSC-HDTV and CCDC), which also used this progressive camera. The somewhat higher noise level of this camera accounts for the observations that the 720-P format of the GA system was judged generally inferior to the 1080-I format in this experiment, in contrast to the results reported above. Other tests in this round used transconverted 1035-I material as a source for 720-P video, which provided both formats with relatively low-noise source material but produced interlace artifacts on the 720-P format.

The Task Force observers ranked the reproduced images on a 10-point scale (10 being best). There was no forced ordering (*i.e.*, ties were permitted). The 10-point scale also allowed the observers to indicate large relative differences in performance, where appropriate. In order to establish the range of image quality they would be considering, the observers pre-viewed some of the images without ranking them.

This test was performed independently of the Free Form Viewing Test (#2.1). Any differences in image ratings are either a consequence of the different contexts of the two tests or of the different image sources explained above.

Interpretations of the Data:

The attached Table 2-3 show the scores for each system on each image obtained by averaging the scoring of the five observers.

Identical scores do <u>not</u> mean that the images and their artifacts were indistinguishable. Identical scores mean that the overall image quality (the net effects of all visible artifacts) was judged the same. Sometimes this involved trade-off among the subjective impressions of different artifacts (*e.g.*, interlace artifacts versus quantization noise, pulsating noise vs. more random "white" noise, etc.).

Conclusions:

Both progressive and interlaced modes of the GA system were judged clearly superior to any and all of the previous systems.

Detailed Test Results:

The images and relative scores are presented in the following table.

Table 2-3

Image	Ref 1035I	GA 1080I	GA 720P	DigiCipher 960I	DSC-HDTV 720P	AD-HDTV 960I	CCDC 720P
Metal Table & Chairs	9.3	9.2	8.9	8.0	5.2	7.8	7.3
Vines	9.0	9.0	8.6	7.7	4.8	8.2	6.4
Wavy Wall	9.3	9.3	9.1	8.3	7.9	8.8	7.0
Tulips	9.3	8.6	8.2	7.8	7.9	7.1	6.6
Sculptures	9.3	8.8	8.5	8.3	6.9	7.4	6.4
Fruits & Vegetables	9.4	9.2	9.4	8.8	8.1	7.5	7.3
Toys	8.7	8.2	7.6	7.5	6.8	6.8	5.5
Girl with Toys	9.0	8.5	7.7	7.9	7.1	7.0	6.1
Memorial Arch	9.5	9.4	8.8	8.1	7.5	7.7	7.5
Woman with Roses	9.5	8.9	8.6	8.3	7.6	8.4	6.8
Cheshire Cat	9.3	9.1	8.7	8.0	7.0	7.5	7.1
Window	9.4	8.8	7.4	8.5	6.8	8.6	6.5
Fax Machine	9.3	8.3	7.6	7.5	7.5	8.1	6.6
Mannequins	9.1	8.1	6.9	8.0	7.1	7.6	5.6
Living Room	8.7	8.0	6.7	7.2	6.4	7.2	5.6
Den	8.9	8.5	7.0	8.0	6.9	8.2	6.0
Park Ride	8.9	8.6	7.5	8.1	7.2	8.0	6.4
Audience	8.7	8.3	6.9	8.1	6.8	8.1	6.4
Woman and Room	9.3	8.3	7.3	7.9	6.9	8.2	6.4
Lamp	8.5	8.2	7.3	8.2	6.9	7.9	6.6
Co-Channel	9.2	8.7	7.3	8.4	6.6	8.5	6.5

Table 2-3 (continued)

Image	Ref 1035I	GA 1080I	GA 720P	DigiCipher 960I	DSC-HDTV 720P	AD-HDTV 960I	CCDC 720P
Rotating Pyramids	9.4	7.5	8.6	6.0	4.3	5.4	6.5
Carousel 24 fps film	7.8	7.6	6.4	7.6	6.4	7.6	6.1
Bridge 24 fps film	7.4	7.2	6.4	7.2	5.8	7.2	6.1
Bridge 30 fps film	7.8	7.5	6.6	7.4	6.2	7.2	6.2
Helicopter	8.7	8.4	6.9	8.0	6.2	8.2	5.6
Slinky	8.7	8.2	8.7	7.6	6.4	7.6	6.2
Tube	8.9	8.7	9.3	8.5	8.2	8.3	7.9
Skiers	8.6	8.4	8.4	7.9	8.1	7.7	7.7
School Yard	8.4	7.9	7.3	7.4	7.0	7.5	6.4
Amusement Park	8.2	7.9	8.3	7.7	7.4	7.6	7.6
Turbo	8.6	7.9	7.9	7.5	6.9	6.4	6.6

Part III

Record of Test Results

for

digital HDTV
Grand Alliance System

from

Video Subjective Tests

Conducted by

Advanced Television Evaluation Laboratory
(May 29 - August 16, 1995)

Advanced Television Evaluation Laboratory

Centre de Recherches sur les communications
3701 Carling Avenue, P.O. Box 11490, Station H • Ottawa, Ontario K2H 8S2 CANADA • 613/991-3313